On the Nature of Soil Moisture in Land Surface Models

RANDAL D. KOSTER,* ZHICHANG GUO,+ RONGQIAN YANG,# PAUL A. DIRMEYER,+ KENNETH MITCHELL,# AND MICHAEL J. PUMA@

* *Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland*
\+ *Center for Ocean–Land–Atmosphere Studies, Calverton, Maryland*
\# *National Centers for Environmental Prediction, Camp Springs, Maryland*
@ *Columbia University, New York, New York*

(Manuscript received 9 September 2008, in final form 24 February 2009)

ABSTRACT

The soil moisture state simulated by a land surface model is a highly model-dependent quantity, meaning that the direct transfer of one model's soil moisture into another can lead to a fundamental, and potentially detrimental, inconsistency. This is first illustrated with two recent examples, one from the National Centers for Environmental Prediction (NCEP) involving seasonal precipitation forecasting and another from the realm of ecological modeling. The issue is then further addressed through a quantitative analysis of soil moisture contents produced as part of a global offline simulation experiment in which a number of land surface models were driven with the same atmospheric forcing fields. These latter comparisons clearly demonstrate, on a global scale, the degree to which model-simulated soil moisture variables differ from each other and that these differences extend beyond those associated with model-specific layer thicknesses or soil texture. The offline comparisons also show, however, that once the climatological statistics of each model's soil moisture variable are accounted for (here, through a simple scaling using the first two moments), the different land models tend to produce very similar information on temporal soil moisture variability in most parts of the world. This common information can perhaps be used as the basis for successful mappings between the soil moisture variables in different land models.

1. Introduction

Most land surface models (LSMs) used with atmospheric general circulation models (AGCMs) keep track of the moisture state of the soil with a prognostic (state) variable called "soil moisture" or "soil wetness," typically defined at a number of vertical subsurface levels. Given the general dearth of in situ large-scale observations of soil moisture, this model-generated quantity is often made available to the scientific community as a data product. The National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis, for example, provides global "soil wetness" values on a T62 Gaussian grid (Kalnay et al. 1996), and the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) provides volumetric soil water in four soil layers on a roughly 1° spatial grid (Simmons and Gibson 2000). The Japanese 25-yr reanalysis (JRA-25) includes a dimensionless global data product called "soil wetness" (http://jra.kishou.go.jp/JRA-25/elements_en.html). The Global Land Data Assimilation System (GLDAS) provides real-time estimates of soil moisture fields derived from land model integrations of realistic atmospheric forcing (Rodell et al. 2004).

The scientific community, in turn, is filled with potential users of soil moisture data. Forecasts of weather and seasonal climate variations in particular can benefit from realistic initial states for the soil moisture variables used in the forecast system (Fennessey and Shukla 1999; Douville and Chauvin 2000; Koster et al. 2004). Models of ecosystem and carbon cycle processes (Yuste et al. 2007; Friend and Kiang 2005) require soil moisture estimates as inputs since soil moisture regulates both soil respiration and plant water stress, which affects stomatal conductance and carbon uptake. Agricultural (crop growth) models similarly rely on soil moisture estimates (Rosenzweig et al. 2002). Soil moisture estimates can benefit the modeling of dust generation (Fecan

Corresponding author address: Randal D. Koster, Global Modeling and Assimilation Office, Code 610.1, NASA GSFC, Greenbelt, MD 20771.
E-mail: randal.d.koster@nasa.gov

DOI: 10.1175/2009JCLI2832.1

© 2009 American Meteorological Society

et al. 1999) and trace gas fluxes (Holtgrieve et al. 2006) from the earth's surface. Groundwater resources can be estimated from Gravity Recovery and Climate Experiment (GRACE) satellite data, but only if independent near-surface soil moisture estimates are available to subtract from the total terrestrial water signal (Rodell et al. 2007). Beyond the immediate scientific community, potential users of soil moisture data include those interested in the pricing of agricultural futures and those monitoring agricultural drought (Narasimhan and Srinivasan 2005).

These are but a few of many potential users of soil moisture data. A question, then, naturally arises: are the reanalysis and LDAS-type soil moisture products being provided to users in a form appropriate to these applications? At first glance, one might think that they are. The products are generally given straightforward names (''soil moisture,'' ''soil water,'' ''soil wetness'') with seemingly unambiguous meanings and clear-cut units (e.g., percent volumetric water content). Uninitiated users might reasonably assume that, once model-specific treatments of such parameters as soil layer depth, soil texture, and porosity are accounted for, reanalysis-generated soil moisture products are ready to be used directly in their own specialized models.

This, however, is not the case, as indicated to varying degrees by a wealth of literature (e.g., Koster and Milly 1997; Chen et al. 1997; Entin et al. 1999; Schaake et al. 2004; Dirmeyer 2004). Simulated ''soil moisture'' does not have an unambiguous meaning. It is a strongly model-specific quantity, essentially an ''index'' of the moisture state, with a dynamic range defined by the specific evaporation and runoff formulations utilized by the given model (Koster and Milly 1997), *in addition to* model-specific soil parameters such as porosity, hydraulic conductivity, wilting point, and layer depth. Large differences are seen in the soil moisture products generated by different land models, even when the models are driven with precisely the same meteorological forcing (Dirmeyer et al. 2006). Given these differences, the direct transfer of one model's soil moisture into another model is, in fact, inappropriate.

While this issue is generally well understood by the land surface modeling community, it is often overlooked or unappreciated by the scientific community at large, that is, by many potential users of the data. Over the years, the authors of this paper have come across numerous requests from across the community for soil moisture products, and they have seen numerous proposals for inappropriate uses of these data, uses that neglect the model dependencies discussed here. While we avoid here the naming of investigators and projects that have made this mistake, we assert that the problem is widespread enough to justify a highly focused study on the topic, one that will make clear the potential problems associated with transferring soil moisture contents between models and will point to a potential resolution of such problems: the proper statistical interpretation of model-generated soil moisture data.

We present such a study here. We begin in section 2 with two examples that illustrate clearly the potential danger of utilizing a reanalysis-based soil moisture product directly within another model. The discussion following section 2 has two purposes. First, in section 4, we highlight with a single, focused, global-scale study (outlined briefly in section 3) the degree to which soil moisture values produced by different land surface models do indeed differ, even when the models are driven with identical atmospheric forcing. Then, in sections 5 and 6, we show that the different soil moisture products, when properly interpreted, nevertheless are mutually consistent and thus do have intrinsic value. These results support the suggestion of Saleem and Salvucci (2002) that the true information content of soil moisture data lies not necessarily in their absolute magnitudes but in their time variability. In essence, we demonstrate here on the global scale that, if the nature of soil moisture in two models is well understood, a ''mapping'' could be derived that would allow the soil moisture variable from one model to be transformed and then used effectively in the second. We conclude in section 7 with some general remarks about model-generated soil moisture data and its applications.

2. Consequences of inappropriate transfer of soil moisture data

a. Seasonal forecasting example

A demonstration of the practical dangers of directly using one model's soil moisture within another is provided by a recent example from NCEP. Seasonal forecasts with the NCEP coupled global Climate Forecast System (CFS) require initialization of ocean, land, and atmosphere states. In developing its next-generation CFS, NCEP upgraded to a substantially newer land model (the Noah model), while retaining the option in the CFS to execute the previous, older land model [the Oregon State University (OSU) model]. In early pilot experiments with the new CFS, experiments were carried out with both the new Noah model and the old OSU model.

At the time of these early experiments, neither a coupled global reanalysis nor an uncoupled land-only global reanalysis using the newer land model was available. Hence, when the new CFS was run with either the new

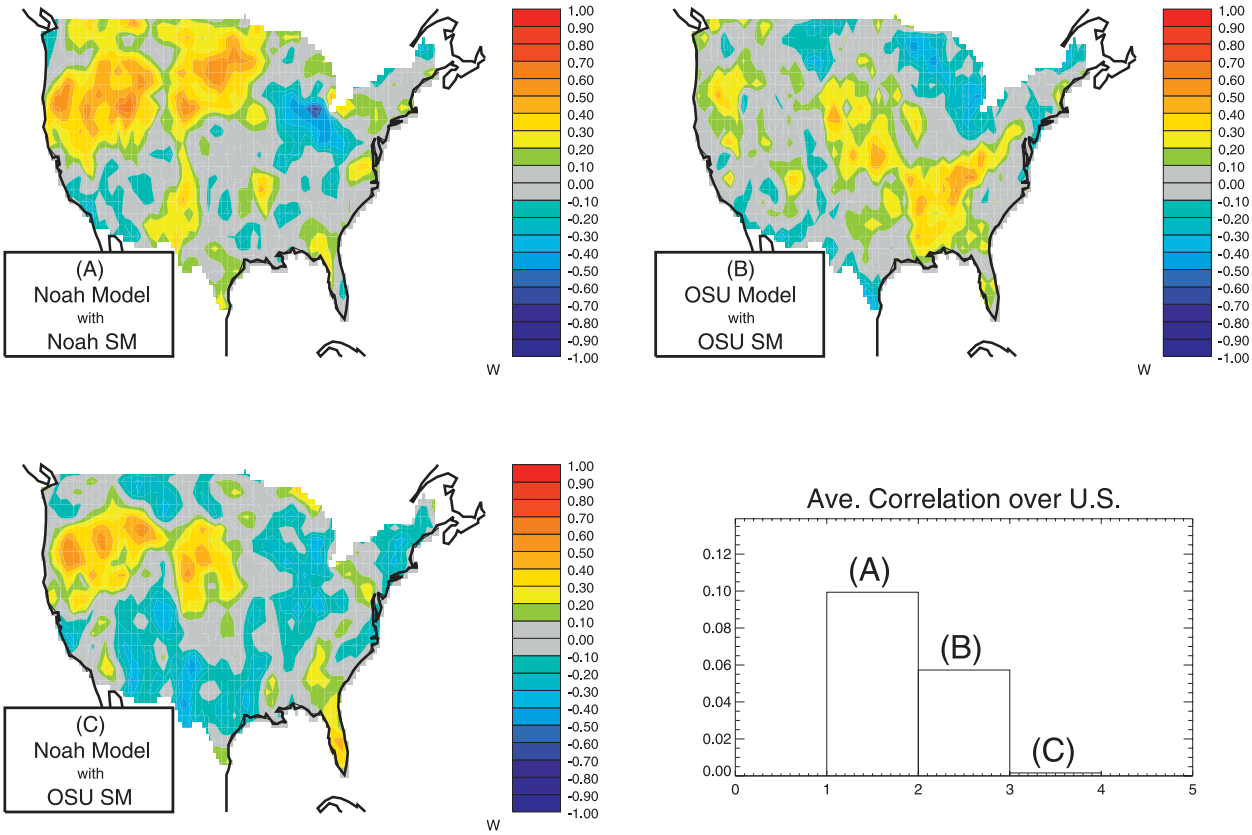

FIG. 1. Skill of seasonal precipitation forecasts (as measured by the correlation coefficient between the forecasted quantities and observations) achieved with different combinations of forecast systems and soil moisture initial conditions: (a) CFS forecast system using the Noah land model, initialized with land states from an offline analysis with the Noah land model; (b) same CFS forecast system but using the OSU land model, initialized with land states from the NCEP/DOE global reanalysis that used the OSU land model; and (c) same CFS forecast system using Noah land model, but initialized with land states from the NCEP/DOE global reanalysis that used the OSU land model. (bottom right) Spatial averages of the correlations produced over the continental United States for each of the three combinations in (a)–(c).

Noah land model or the old OSU land model, initialization of land states came from the NCEP/Department of Energy (DOE) global reanalysis (Kanamitsu et al. 2002), which used the old OSU land model. Thus the CFS–OSU configuration was initialized with model-appropriate soil moisture states, whereas the CFS–Noah configuration used model-inappropriate, OSU-based values (though suitably adjusted to account for changes in model soil layer thicknesses). Concerns about the appropriateness of the CFS–Noah initialization helped spur the execution of an offline, uncoupled global land reanalysis using the Noah model that spanned the period 1979–2006. The Noah-based land states produced in this offline land reanalysis then served as (model appropriate) initial land states for a second set of CFS–Noah experiments.

Specifically, the CFS experiments reported here consist of a series of summer season [June–August (JJA)] forecasts using a CFS forecast system under three sets of land-model/land-initialization combinations: (i) the OSU land surface model initialized with land states from the NCEP/DOE global reanalysis using the OSU land model; (ii) the Noah land model, also initialized with OSU model land states from the NCEP/DOE global reanalysis (representing the newer system, with inconsistent land initialization); and (iii) the Noah land model initialized with the independent global offline (land only) reanalysis with the Noah land model (representing a targeted new, and more appropriate, system). Each combination produced 10-member ensemble forecasts for JJA for each of the 25 years from 1980 to 2004, based on 10 initial times spanning late April to early May. All forecasts utilized the same model resolution and the same sets of atmospheric and oceanic initial conditions.

Results are shown in Figs. 1 and 2 . The first three panels in each figure show, for each of three model/initialization combinations, the skill of the forecast system in predicting seasonal (JJA) averages of precipitation and air

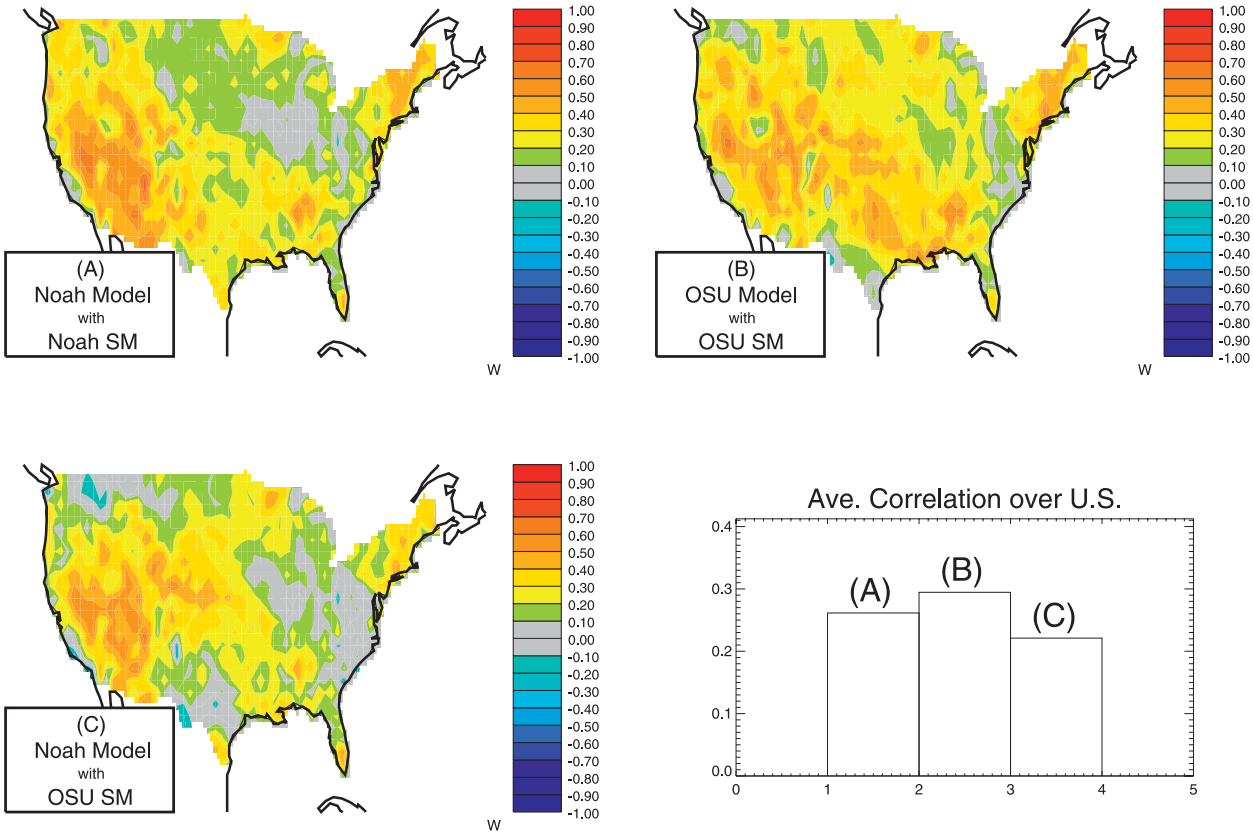

FIG. 2. As in Fig. 1, but for 2-m air temperature.

temperature. Skill here is measured as the temporal correlation between the seasonal mean of the forecasted variable and the corresponding observations across the 25 yr period. [For precipitation the observations come from the Climate Prediction Center gauge-only precipitation analysis (Higgins et al. 2000), and for air temperature they come from the Global Historical Climate Network/Climate Anomaly Monitoring System global air temperature analysis (Fan and van den Dool 2008).] CFS is run here at T126 horizontal (spectral) resolution, with 64 vertical sigma layers and with the atmospheric physics of the NCEP operational medium-range Global Forecast System (GFS) as of June 2007. The ocean model used is the Modular Ocean Model version 3 (MOM3) from the Geophysical Fluid Dynamics Laboratory.

For two of the panels in Figs. 1 and 2, the initialization of prognostic land states (including soil moisture, the key land variable providing seasonal predictability) is appropriate for the particular land model used in the forecasts. The third panel, however, shows the results obtained with the Noah model initialized with OSU model-based states. A comparison of the skill levels—and particularly of the continental U.S. averages shown in the final histograms of Figs. 1 and 2—shows that the two cases using model-consistent initial soil moistures perform best. For precipitation, *t* tests were performed on the U.S. average anomaly correlations. The results show that, statistically speaking (at the 90% level, assuming 24 degrees of freedom), the skill of the two cases with model-consistent initialization are indistinguishable from each other, whereas both are superior to the case with inconsistent initialization.

The OSU and Noah soil moisture products are, indeed, systematically different (not shown), with the OSU soil moistures consistently lower in most global land locations and at most times of the year. Figures 1 and 2, along with the *t*-test analysis, show that direct insertion of OSU soil moistures into the Noah model leads to a degradation in forecast skill. The results of this experiment have helped motivate an even more recent updating of the new NCEP seasonal forecast system, one that uses a global, offline analysis with the Noah land model to initialize the Noah land model component of the system. At the time of this writing, the new system is in operational production of 30-yr reforecasts spanning 1979–2008 in preparation for real-time operational use.

b. Ecological modeling example

The modeling of dynamic vegetation and associated carbon budgets is an area of active research and development (e.g., Moorcroft 2003; Thornton and Zimmermann 2007). One recent effort, based at the NASA Goddard Institute for Space Studies (GISS), focuses on the development of a Dynamic Global Terrestrial Ecosystem Model (DGTEM) that uses a unique representation of mixed vegetation canopies to characterize water, energy, carbon, and nitrogen budgets within a vegetation system; seasonal and interannual variations in the growth and senescence of vegetation; and long-term (out to centuries) vegetation change. The DGTEM, labeled ''Ent,'' is modular in design and is intended for use with earth system models, notably those run by GISS and by the NASA Global Modeling and Assimilation Office (GMAO). Key to the example examined here is the Ent DGTEM's treatment of soil respiration, that is, the release of carbon dioxide from the soil through the decay of biological material. Ent's soil biogeochemistry module is based largely on the Carnegie Ames Stanford Approach (CASA) model (e.g., Randerson et al. 1997), but with modified relationships between soil moisture, temperature, and respiration rate, as derived from the reanalysis of field data measured by Del Grosso et al. (2005).

The comparisons presented here focus on Ent simulations of soil respiration at the Vaira Ranch site in California in 2002 under three prescribed time series of volumetric soil moisture for the top 30 cm of soil: (i) soil moistures that were measured at the site, (ii) soil moistures produced by the NCEP–NCAR reanalysis at the grid cell containing the site (aggregated appropriately across the topmost model layers), and (iii) the corresponding soil moistures produced by ERA-40. We first show, in the top panel of Fig. 3, the differences in these time series. (Note that, although monthly values are shown, the temporal resolution of the raw data was daily or higher.) All three time series show higher soil moisture values in winter and lower values in summer and fall, in straightforward agreement with the winter-based seasonal cycle of California precipitation. Nevertheless, they also show distinct differences. The observed series dips to low values by April, and the soil remains dry through August. The reanalysis-based moistures show a slower, steadier decline over the course of the spring and summer. Furthermore, the seasonal amplitude of the signal is much larger for the NCEP–NCAR reanalysis, at least twice that of the ECMWF reanalysis.

As suggested above in section 1 and as demonstrated below in section 4, the models' soil moisture variables are best thought of as ''indices'' of moisture content and

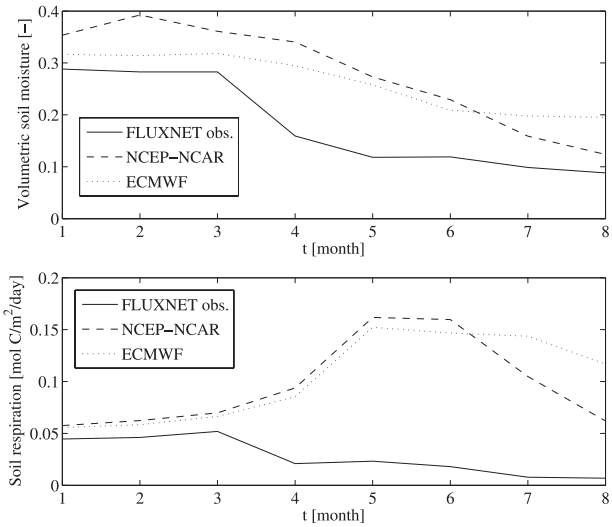

FIG. 3. (top) Comparison of soil moisture measured at the Vaira Ranch FLUXNET site in California with the soil moisture in the corresponding grid cell produced by the NCEP–NCAR reanalysis and ERA-40, expressed as dimensionless volumetric soil moisture. (bottom) Soil moisture respiration (moles of carbon m^{-2} day^{-1}) produced by an ecological model using the three soil moisture time series as inputs.

do not correspond to measurable soil moisture amounts; for this reason alone, we should expect at the outset that the three time series would not agree. A second explanation, however, should also be considered. The observed time series was extracted at a point, whereas the model-based time series represent large gridcell averages. It is quite possible that spring rainfall at the site was significantly lower (or perhaps evaporation or runoff was significantly higher) than the large-area mean, thereby accounting for the faster drop in the observed soil moisture. Even if this were true though, the top panel would still demonstrate problems with assuming that reanalysis-based soil moistures can be applied with confidence to point models. Furthermore, this ''point-scale'' reasoning does not explain the differences in the two reanalysis soil moistures; if anything, under this argument the higher-resolution ECMWF data should show the larger amplitude.

The bottom panel in Fig. 3 shows the soil respiration rates determined by the Ent DGTEM under the three different soil moisture time series. (All other forcings in the simulations were identical.) The reanalysis-based soil moistures produce respiration rates roughly five times higher than do the observed soil moistures during spring and summer. The two reanalysis products produce similar respiration rates until July, when the rates for the ECMWF soil moisture begin to exceed significantly those produced with the NCEP–NCAR soil moisture, by about a factor of 2 in August. This example thus

demonstrates that, if accurate estimates of soil respiration were needed for, say, a carbon budget analysis, the direct use of reanalysis soil moisture products in a respiration model could lead to substantial error.

3. The Global Soil Wetness Project

Both of the examples above demonstrate the dangers associated with treating a reanalysis soil moisture as one would treat, say, a model or reanalysis air temperature—that is, as a quantity that is intrinsically comparable (in magnitude and range) to a true, observable geophysical quantity. With these examples in mind, we now turn our attention to a study performed with data from the second phase of the Global Soil Wetness Project (GSWP-2). GSWP-2 was a component of the modeling efforts of the Global Energy and Water Cycle Experiment (GEWEX), efforts focused on the large-scale performance and assessment of LSMs commonly used in weather and climate models and on the production of global datasets for use by the earth science community (Dirmeyer et al. 2006). In GSWP-2, a number of LSMs were run on a common, global $1° \times 1°$ grid (excluding Antarctica), each driven by the same 10-yr time series of observations-based meteorology. The forcing time series were derived from a combination of gridded 3-hourly reanalysis output, 3-hourly satellite-derived surface radiation data, and monthly gridded observational data (Zhao and Dirmeyer 2003). The precipitation data, in particular, were forced to agree with observational datasets at the monthly time scale. The core period of 1986–95 covered by the offline simulations corresponds exactly to that of the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II dataset (Hall et al. 2006).

As part of GSWP-2, each participating LSM produced a standard set of output grids at the same 1° resolution, data comprised of land surface state variables, fluxes, and other diagnostic quantities. Among these are several variables related to soil moisture, including the change in column-integrated water content, volumetric soil moisture in each model layer (layers vary amongst the models), and total column soil moisture. In this paper, we focus on the average soil moisture in the top meter of soil (nominally, the ''root zone''), as computed from layer totals using an objective procedure (Guo et al. 2006). We normalize this quantity by dividing it by the total water holding capacity of the top meter (porosity times depth), yielding a dimensionless quantity known as the degree of saturation, which can vary from 0 to 1.

Note that until now in this paper, we have freely utilized the term ''soil moisture'' in deference to its widespread usage in the community, despite recognizing that it is a model-specific quantity with no direct observational analog. Even ''degree of saturation'' has a precise hydrologic meaning (e.g., Eagleson 1970), one inappropriate for such a model-specific variable. Thus, for greater precision in our discussion of the GSWP-2 analysis, we will hereafter refer to the computed degree of saturation for a GSWP-2 model as a root-zone ''soil wetness index.'' Always keep in mind, however, that, given its simple definition, all stated results regarding ''soil wetness index'' apply directly to model-generated ''soil moisture'' or ''soil wetness'' values, of the type provided as data products from reanalyses or LDAS systems.

Data from 7 of the 15 participating models are used hereafter in this study. These 7, which can be considered representative and which for simplicity are referred to in this paper as models A–G, were selected from the 15 based on the availability of water holding capacity data and an independence of historical development (i.e., data from different versions of the same model, though available through GSWP-2, were not considered, though experience tells us that even different versions of the same model can produce the types of differences illustrated below).

For a proper statistical analysis, the time series examined should be stationary; long-term trends need to be removed. Thus, prior to all analyses, the time series of soil wetness index produced by each model at each grid point were ''detrended'' by first fitting a line, through simple regression, through each time series, translating the line vertically downward to give it zero mean over the 10-yr period, and then subtracting this line from the time series. This detrending was, in fact, mostly performed to prevent spurious intermodel time series correlations in desert regions associated with incomplete spinup—if two models are continuously drying over the 10-yr period, the two raw (not detrended) time series will show a strong correlation that has little meaning. The detrending has no noticeable impact on our results outside of desert regions.

4. Examples of soil wetness index inconsistency

Figure 4 shows, for five representative sites, the 10-yr time series of daily averaged root-zone soil wetness index produced by the seven GSWP land models. Again, the wetness index can vary by definition from 0, for completely dry conditions, to 1, for completely saturated conditions.

The model differences are substantial. The examples, in agreement with the earlier findings of Chen et al. (1997) and Entin et al. (1999), illustrate the wide variety of wetness index behavior inherent in different state-of-the-art land surface models, both in terms of mean

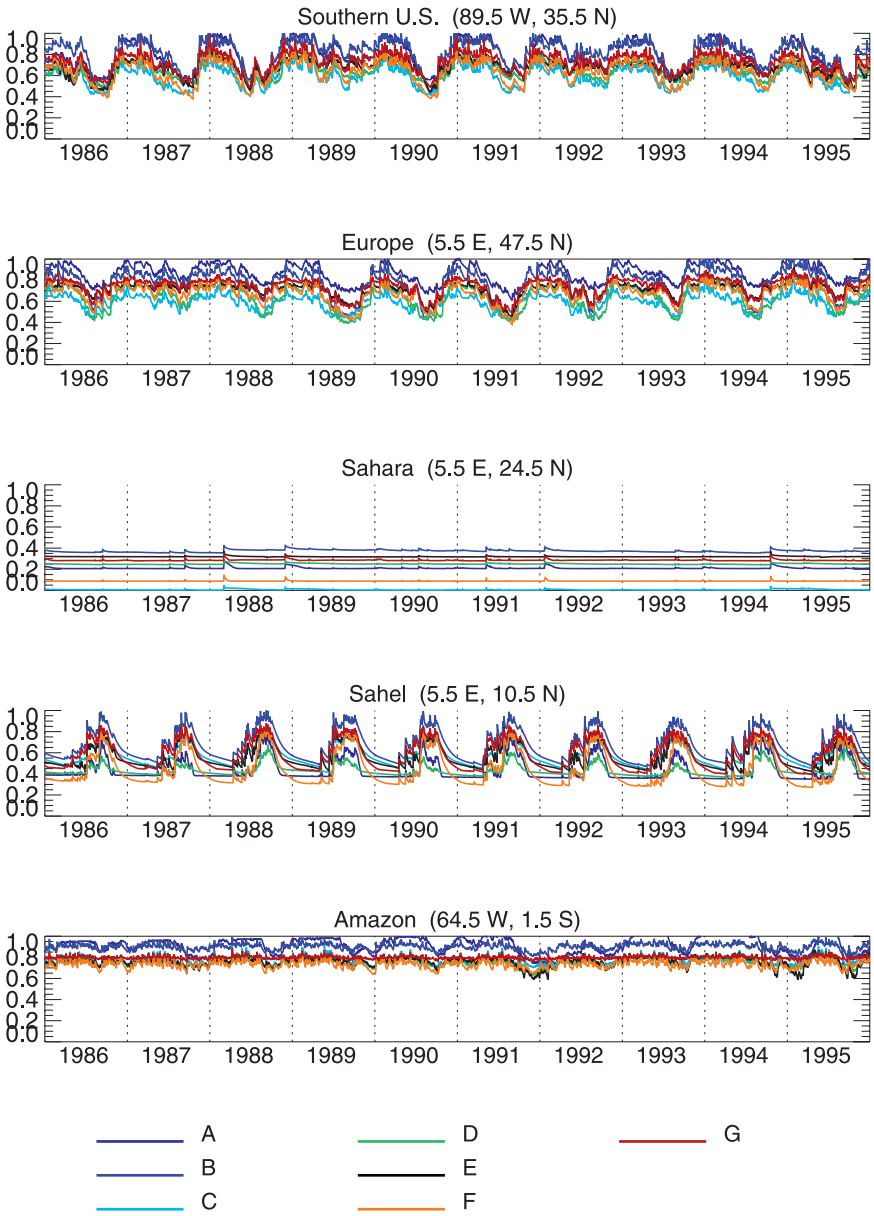

FIG. 4. Time series of root-zone soil moisture (degree of saturation) for each of the seven considered GSWP-2 models at five representative locations.

climatology (models C and D can be drier than model A throughout the simulation period) and dynamic range (in the Sahel the range for model D is about 0.2, whereas that for model B is about 0.5). The figure shows clearly that the wetness index variable, and thus the ''soil moisture'' variable from which it was derived, in a given model is specific to that model. Naturally, it would behoove climate modelers to characterize fully the values of the variable produced by their land model prior to performing sensitivity studies involving that variable. In examining, for example, wet and dry soil impacts on weather or climate, the assigned ''wet condition'' or ''dry condition'' should be made sensibly, in the context of the model's own dynamic range, as determined through analyses such as that behind Fig. 4. Setting the initial perturbations too large, outside this inherent range, could seriously reduce the usefulness of such a study.

More to the point, the intermodel wetness index differences in Fig. 4 illustrate the potential problems associated with the direct replacement of one model's soil moisture with that from another. In the Amazon, for

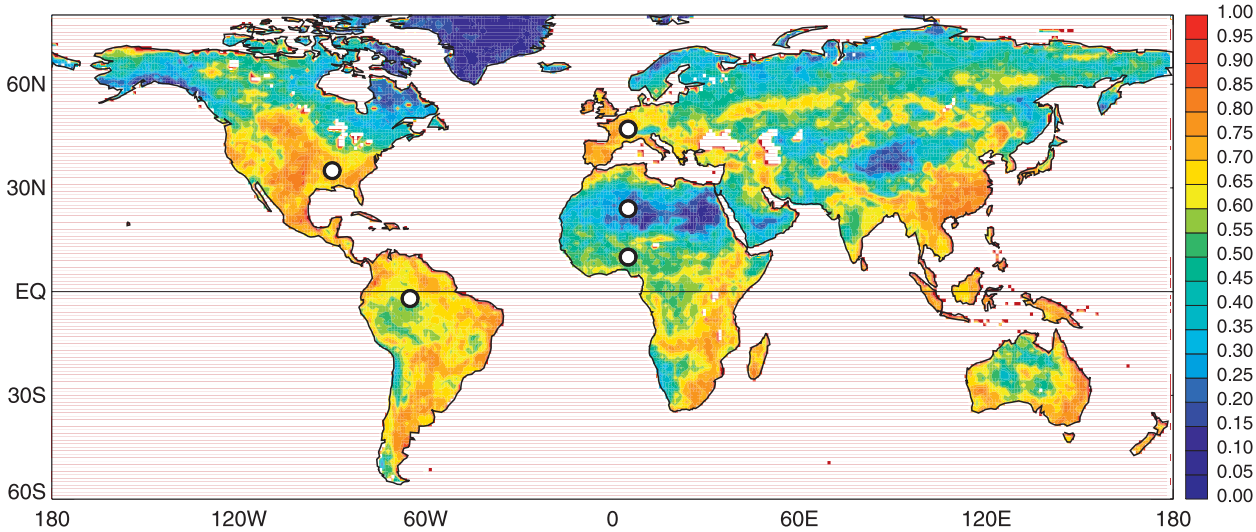

FIG. 5. Map showing the average r^2 values (see text) computed from the LSM soil moisture products, a measure of the degree to which the LSM products contain the same information on the time variability of soil moisture. The five white circles identify the gridcell locations examined in Figs. 4 and 6.

example, the wettest state for model F is about equal to the driest state for model B. As a result, regardless of how wet the soil *should* be (following, say, an extensive rainy period), model B would always be initialized very dry if it used model F's soil moisture without modification. Intuitively, using model F's products to initialize model B would significantly degrade the performance of model B, even if the model F product was part of a trusted atmospheric reanalysis or offline land analysis. Again, such considerations lie behind the degradation of skill seen in Figs. 1 and 2, for the OSU land model tends to produce drier soil moistures than the Noah land model under the same atmospheric forcing.

5. Underlying information content on temporal soil moisture variability

We emphasize again that an LSM's soil moisture variable is best thought of as a model-specific index of wetness, a monotonic function of the true average soil moisture that it is trying to represent. Because different LSMs use, in effect, different monotonic functions, the "soil moistures" they produce have different climatological means and variability characteristics. Even so, the time variations of the different LSM products should be correlated; for example, regardless of the LSM used, an excessively rainy period should lead to a higher value for its soil moisture variable than a dry period. Of course, perfect correlations will never be achieved, if only because soil moisture, through the conservation equation, also depends in part on runoff and evaporation fluxes, and different LSMs use different formulations for these fluxes.

We characterize here the "level of agreement" between different soil moisture products using such temporal correlations. (Again, throughout this paper, we focus on the total moisture in the root zone; a supplemental analysis of total column soil moisture produced equivalent results.) At each GSWP-2 grid cell, we first compute the square of the correlation between the time series of wetness index produced by each independent pairing of LSMs. The daily (detrended) data are used for this calculation with the mean seasonal cycle removed. For seven LSMs, we obtain 21 r^2 values. We then compute the average of these 21 values and call this average our measure of LSM agreement.

The global distribution of the average r^2 values is shown in Fig. 5. The average r^2 is low in Eastern Hemisphere deserts and in high northern latitudes. Almost everywhere else, however, LSM agreement is high, with average r^2 values exceeding 0.5, often substantially. The average r^2 values indicate that, despite the known differences in soil wetness index products (e.g., Fig. 4), the LSM products indeed *share* a great deal of underlying information about the time variability of the wetness index. This shared information suggests that a mapping of soil wetness indices, and thus soil moistures, between models may be possible.

Schaake et al. (2004), using a somewhat different analysis approach over the continental United States, produced a similar result; they found that four different models driven with the same meteorological forcing

tended to produce similar temporal variations in wet areas but not in dry areas. In effect, the present study affirms their result using a greater number of models and a larger (global) spatial domain.

The low correlations in the deserts are, in some ways, not surprising. Arguably, model agreement should be largest in regions with large variations in precipitation forcing—a larger precipitation variance suggests a larger variation in moisture storage that all of the models can more easily capture. The deserts, by definition, have low mean precipitation and a correspondingly low precipitation variance. In high northern latitudes, the differences may relate in part to differences in the parameterization of snow and frozen soil.

6. Mapping of soil wetness index, and thus soil moisture, between models

a. Scaling with the first two statistical moments

The key to any proper transferability of soil moisture states between LSMs lies in the recognition of and correction for the differences in the statistical moments of the LSMs' soil moisture distributions. If the differences in the LSM-based soil wetness indices in Fig. 4 relate mostly to differences in their means and standard deviations, then all models would produce roughly the same time series of standard normal deviates w':

$$w' = (w - w_m)/\sigma_w, \tag{1}$$

where w_m is the mean wetness index for the given LSM at the point and time of year in question and σ_w is that LSM's standard deviation of wetness index for that point and time of year. Dirmeyer et al. (2004), using (1), suggest an equation for transforming a wetness index value for model A into one that can be used in model B:

$$w(\mathrm{B}) = [(w(\mathrm{A}) - w_m(\mathrm{A}))/\sigma_w(\mathrm{A})]\,\sigma_w(\mathrm{B}) + w_m(\mathrm{B}). \tag{2}$$

For example, using (2), a relatively dry state in model A would be mapped into a correspondingly dry state in model B.

Of course, such a transformation ignores the fact that soil moisture, and thus wetness index, distributions are often skewed rather than Gaussian, particularly when mean soil moisture is at the drier or wetter end of the spectrum. The non-Gaussian properties could allow, for example, the generation through (2) of scaled soil wetness indices that fall below 0 or exceed 1, indices that would then have to be constrained after the fact to sensible values. In principle, we could address this and other problems associated with the underlying non-Gaussian distributions explicitly using a transformation method that involves more sophisticated statistical techniques, such as cumulative distribution function (CDF) matching (e.g., Wood et al. 2002). Given, however, the availability of only 10 years of data from GSWP-2, use of such techniques, while not impossible (see, e.g., the ergodic approach of Reichle and Koster 2004), is nevertheless difficult, and thus for convenience we focus in this paper on the use of the simpler (2). Our justification for going ahead with (2) despite its limitations is simple: both the non-Gaussian nature of soil moisture and, indeed, our inability to estimate w_m and σ_w accurately from only 10 truly independent data points for a given time of year will only hinder our ability to develop an effective mapping between models. Thus, any success in mapping that we obtain through (2) can be considered conservative. As will be seen, our mapping success using (2) despite being "underestimated" is, indeed, considerable.

b. Time series of soil moisture at selected cells

We now (arbitrarily) choose model F as the "target" model. We use (2) to transform the soil wetness index time series of each of the other models to values relevant to model F. For each model, the w_m and σ_w values are computed at each grid cell and each day of the year from the available GSWP-2 data. We use a 31-day moving window, with 15 days on either side of the day in question in each of the 10 years, to provide 310 data points for the statistics; note, however, that, given the long-term memory of soil moisture, the 310 data points are far from independent. Given that soil moisture memory time scales can easily span a month (e.g., Entin et al. 1999), we can at best claim to use 10 fully independent data values to compute w_m and σ_w.

Figure 6 shows, for the five grid cells examined in Fig. 4, the time series of the transformed wetness indices. The time series for the different models are largely coincident at the five grid cells considered. Again, the grid cells shown are representative: by and large, despite the model differences illustrated in Fig. 4 and in accordance with the levels of agreement plotted in Fig. 5, the models tend to contain the same information about the time variability of soil "wetness," information that can be transferred to a specific model (here, model F) using the scaling in (2).

c. Global picture of mapping success

The top panel of Fig. 7 gives a broad-brush indication of where the original (unscaled) model-derived wetness indices, and thus soil moisture values, tend to be different. For each of the 3652 days of simulation at each grid cell across the globe, we computed the intermodel

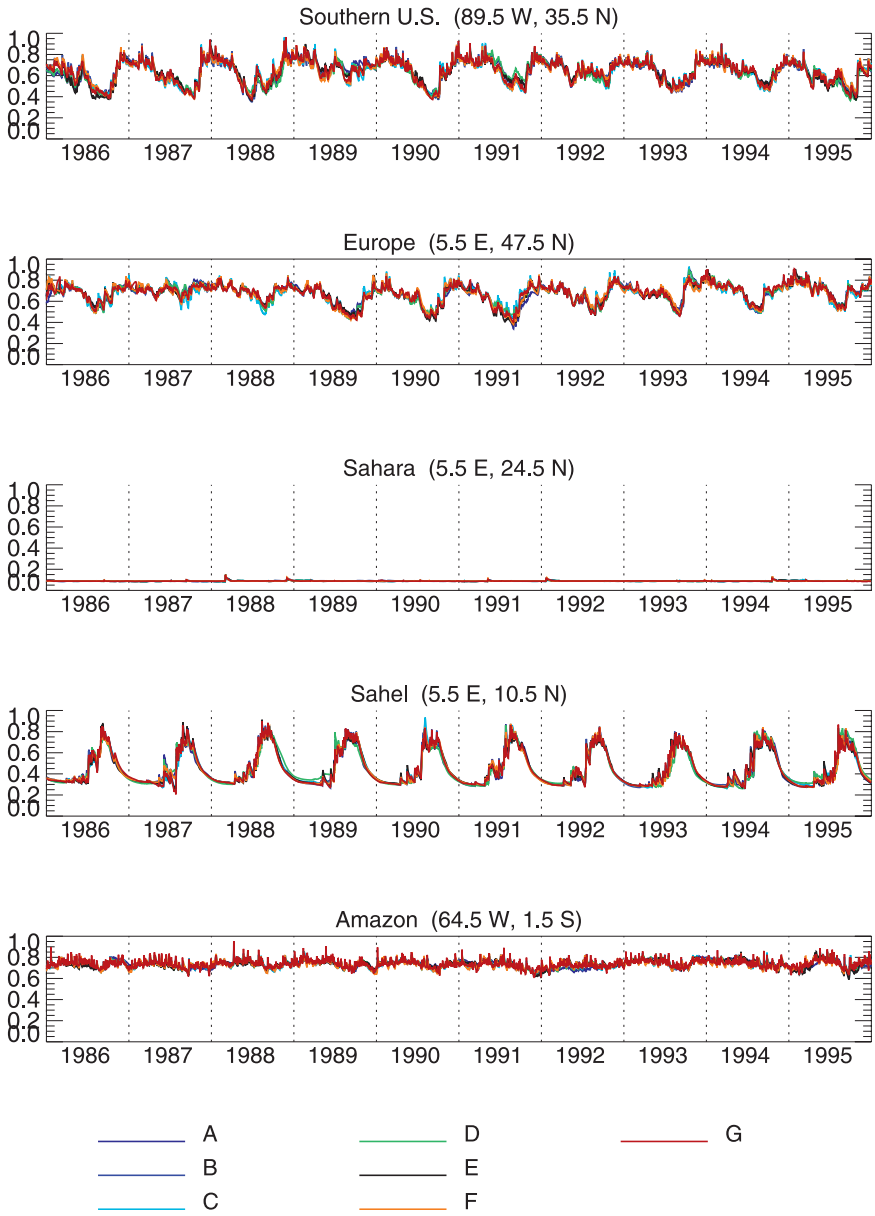

FIG. 6. As in Fig. 4, but for soil moistures scaled with (2) to produce values relevant to model F.

variance of soil wetness index and then averaged these variances over time and plotted the global distribution of their square roots, showing in essence a mean standard deviation field. The standard deviations of the degree of saturation are largest in the Eastern Hemisphere deserts and in far northern latitudes. Across the globe, the standard deviations are almost always above 0.1. Note that, even where the standard deviations are lowest, model differences still prevent a direct transfer of one LSM's wetness index into another. The black circle in South America locates the Amazon grid cell examined in Fig. 4. The standard deviation computed for the grid cell is relatively low, and yet Fig. 4 shows that several models have markedly different wetness index values there.

The bottom panel of Fig. 7 provides the equivalent figure for the transformed soil wetness indices. First, model A was taken to be the "target model," and the wetness indices for the remaining models were transformed with (2) to produce values relevant to model A. The average (over 3652 days) intermodel variances for the transformed indices were computed as above, and their square roots were taken. Next, model B was taken to be the target model, and the process was repeated.

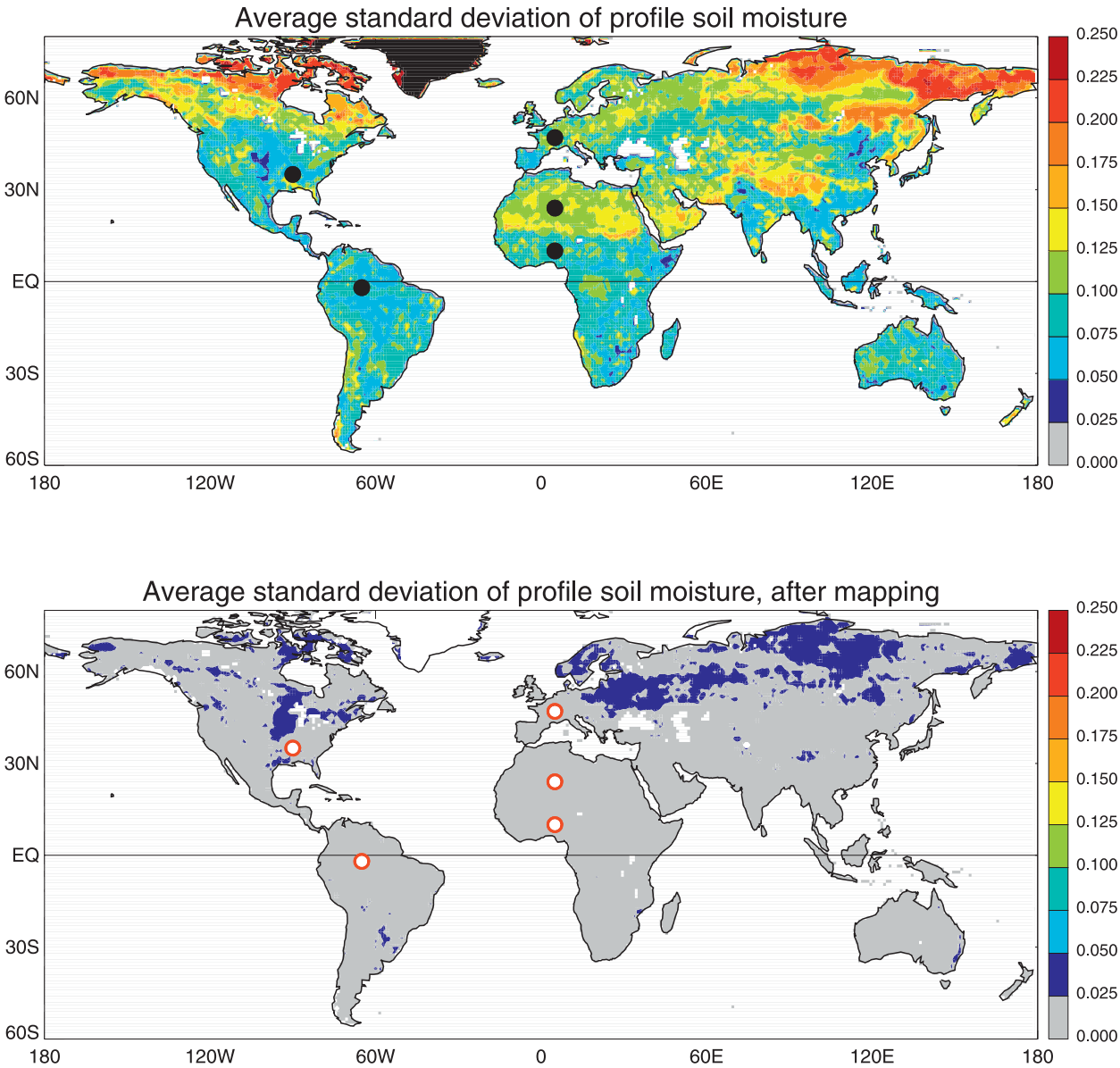

FIG. 7. (top) Map showing the standard deviation across the models of root-zone soil moisture (degree of saturation). The variance was computed for each day of each year and then averaged over the 3652 days of simulation; the square root of the resulting average variance was then found and plotted. The five black circles identify the gridcell locations examined in Figs. 4 and 6. (bottom) As above but for the average root mean variance generated when each model (in turn) is considered the target model and the soil moistures from the other models are scaled accordingly.

After all seven of the GSWP-2 models were used in turn as the target model, we had in hand seven maps corresponding to the top panel of Fig. 7. The bottom panel of Fig. 7 is the average over those seven maps. Clearly, the mapping is successful, bringing all of the LSM products together with an average standard deviation that is almost everywhere less than 0.025. What little standard deviation remains may result in large part from the aforementioned limiting assumptions surrounding the use of (2), namely, the assumptions of normal soil moisture distributions and of our ability to estimate reasonable values of the first two moments from 10 independent data points. Of course, some of it may also relate to true differences in the information content of the models' time series; still, the low standard deviations indicate that such true differences would be small.

Based on Fig. 7, the low average r^2 values found for the desert in Fig. 5 should not be much of an issue when

transferring one LSM's wetness index, and thus soil moisture product, into another. The temporal variability tends to be so low in deserts that, even though the models are not temporally correlated there, the disagreement between variations would be too small to affect a climate simulation. For example, the Saharan point shows a very low average r^2 value in Fig. 5, and yet Fig. 6 shows the time series of scaled soil wetness indices at the point to be almost completely coincident.

7. Discussion

The agreement in the LSM soil wetness index products illustrated in Fig. 5 and the success of the mapping demonstrated in Figs. 6 and 7 echoes somewhat the findings of Dirmeyer et al. (2004). The present study serves as a more focused and controlled analysis of intermodel soil moisture differences since each of the (more abundant) LSMs examined here is forced with identical meteorological data—thus we isolate the impacts of the LSM differences themselves. We similarly affirm and expand on the results of Schaake et al. (2004) by examining a greater number of land models over the full global domain. We find that, while the wetness indices, and thus the soil moisture products, from different land surface models are essentially very different, they contain generally the same information regarding temporal soil moisture variability so that the products generated by one model can potentially be transformed into quantities relevant to another. The presence of similar temporal variability information in the face of distinctly different soil moisture values (in essence a consequence of similarity in the simulated time series of the infiltration, evaporation, and runoff fluxes that control the models' soil moisture states) is a clear demonstration of the ''model-specific index'' nature of the soil moisture variable; it underlines the fact that an LSM-derived ''soil moisture'' is not (as its name implies) a physical quantity that can be directly validated with field measurements.

Is the ambiguity of a given model's soil moisture product acceptable? Or does it imply that modelers, before using their models, need to improve them until simulated soil moisture values are indeed comparable to those of other models and to observations? Arguably, the ambiguity is largely unavoidable given current computational constraints on LSM resolution. The single soil moisture variable at a given vertical level in a land surface model must reflect more than the average soil moisture across, say, 100 km. It must also—somehow—implicitly reflect the spatial variability of soil moisture and the effects of this variability on the surface energy and water fluxes. Even if we could construct a fully reliable macroscale model that represents the effects of porosity, tortuosity, etc., within a local soil volume without resolving these features explicitly (a complex topic in itself to address), soil moisture heterogeneity at larger spatial scales—e.g., at the hillslope scale (tens to hundreds of meters), where water table dynamics may lead to wet conditions at the bottom of a hill and dry conditions near the top—leads to corresponding heterogeneity in evaporation and runoff production (e.g., Famiglietti and Wood 1991). Given such soil moisture heterogeneity, and owing to strong nonlinearities in the governing soil moisture–evaporation relationship that exists at a point, a modeler simply cannot represent the true average soil moisture in an AGCM gridcell area with a single number and use that single number in a locally derived (single site) relationship to obtain the true areally averaged evaporation rate. For realistic evaporation rates so critical to the proper operation of the atmospheric model, either the evaporation–soil moisture relationship must be modified away from the observed relationship at a point (not something land surface modelers claim to know how to do) or the soil moisture variable in the LSM must take on a nebulous meaning, a meaning distinct from ''areally averaged soil moisture.''

Again, although its meaning may be nebulous, we emphasize that the true information content—and thus value—of a model soil moisture product lies not in its *absolute* magnitudes but in its *time variations*. The proper use of a simulated soil moisture product demands an understanding of this basic element of its character.

We can speculate that, as the spatial resolution of land surface models improves and as modelers take better advantage of growing soil moisture databases for the continued improvement of hydrological parameterizations, the disconnect between modeled and observed soil moisture should decrease. Again, though, fine spatial scales of heterogeneity (tens of meters) help shape hydrological processes in nature, and our ability in the foreseeable future to resolve such scales explicitly in global simulations appears limited (though statistical approaches may provide a valid alternative strategy). Again, even with such highly resolved scales, the submillimeter scales of relevance to soil moisture transport within the soil matrix (i.e., the scale of soil pores) will not soon, if ever, be resolved—necessitating in any case a macroscopic representation of soil moisture.

The proper utilization of observations for development, calibration, and validation of models presents additional challenges. First and foremost, in situ soil moisture data have very limited availability, with many large regions of the world still unmonitored. Even where the in situ data exist, their interpretation is often difficult because the measurements are highly localized, and

differences in soil properties can cause important differences in the mean and variance of soil moisture, even over small distances. Variations in forcing (e.g., convective precipitation) add another decorrelating factor on longer spatial scales, though this becomes less important as the time scale increases.

Interestingly, Saleem and Salvucci (2002) scaled each of a number of observational (in situ) soil moisture time series in Illinois with (1) and showed that the ostensibly different time series contain similar information, in direct analogy to our findings above. Saleem and Salvucci indeed suggest that scaling with (1) or with more complex statistical techniques could aid in model intercomparison. Our findings firmly support their suggestion, on a global scale.

The remote sensing of soil moisture via satellite holds particular promise to address the lack of in situ observations across the globe, but much work is still needed to address the proper use of these data for model improvement. In some existing cases, satellite methods seem to return their own forms of soil moisture indices rather than direct soil moisture estimates (Reichle et al. 2007). Also, satellite-based soil moisture data are limited to the top several centimeters (or less) of soil, whereas the soil moisture data of greatest relevance to many climatic applications spans a much larger depth, through at least the root zone of the native vegetation. We note, as an important aside, that satellite-based soil moisture data can indeed contribute to the estimation of root-zone soil moisture through the process of data assimilation (Reichle et al. 2007), essentially the mathematically optimal merging of satellite-based and land-model-based soil moisture estimates. For data assimilation to proceed with the current generation of land surface models, the model-dependent nature of simulated soil moisture must be recognized and accounted for prior to the data merging.

Finally, we note that, while the use of (1) does allow a mapping of the soil moisture variable between land models and while much more reliable approaches for mapping do exist (e.g., CDF matching) and should be used instead if adequate data are available, the use of any mapping comes with a cost—the modeler must generate and analyze a 10-yr (or preferably much longer) time series of soil moistures produced by both land models involved. Generating such time series is conceptually trivial, particularly given the availability of forcing data through GSWP-2, but the production and analysis of the time series may still, for various reasons (development of an offline driver, grid considerations, etc.), pose difficulties for generating a robust mapping algorithm. Model-specific vertical discretizations of the soil column may pose particular, though not insurmountable, difficulties for mapping.

This paper's main result—that the representative cross section of LSM's examined here contain largely the same information on the temporal variation of root-zone soil moisture so that the soil moisture produced by one land model can, *if properly scaled*, be utilized in another—may extend beyond such simple model-to-model mappings. Because an LSM's soil moisture variable represents the time-integrated impacts of antecedent meteorological forcing on the hydrological state of the soil system, it has a potentially profound real-world application: unless all of the models examined here are wrong in precisely the same way, we can speculate from the agreement levels in Figs. 5, 6, and 7 that any of these models could, if forced with realistic meteorological data (and better yet, if also combined with satellite-based soil moisture information via data assimilation), provide a soil moisture product that, when suitably scaled, captures the time variation of real-world soil moisture. Indeed, the impact of any small model dependence that does exist—as represented by the degree to which the r^2 values in Fig. 5 are still less than 1 or as established through other techniques (e.g., Abramowitz and Gupta 2008)—can perhaps be mitigated by using some type of multimodel average. If the forcing data were provided in real time, the state of the soil could be estimated in real time. The model products could, if suitably scaled, serve to monitor the *real world* hydrological state. Perhaps, for example, they could be used to identify such events as the onset and demise of agricultural drought.

A note of caution about this, however, is appropriate here. It is possible that land surface models are, indeed, all "wrong in precisely the same way," reducing their usefulness for such a real-world application. For example, they all lack, to some degree, sophisticated treatments of certain hydrological processes (e.g., baseflow, interflow), and most share similar simple representations of other hydrological processes (e.g., the use of Richard's equation for subsaturated vertical flow). More research is needed into the question of how well land surface models in general capture real-world hydrological processes.

Acknowledgments. This research was supported by funding from the Earth Science Enterprise of NASA headquarters. We gratefully acknowledge WCRP/GEWEX/GLASS for coordinating GSWP-2, and we particularly thank all of the individual modelers who contributed their data to GSWP-2.

REFERENCES

Abramowitz, G., and H. Gupta, 2008: Toward a model space and model independence metric. *Geophys. Res. Lett.,* **35,** L05705, doi:10.1029/2007GL032834.

Chen, T. H., and Coauthors, 1997: Cabauw experimental results from the Project for Intercomparison of Land-Surface Parameterization Schemes. *J. Climate,* **10,** 1194–1215.

Del Grosso, S. J., W. J. Parton, A. R. Mosier, E. A. Holland, E. Pendall, D. S. Schimel, and D. S. Ojima, 2005: Modeling soil CO_2 emissions from ecosystems. *Biogeochemistry,* **73,** 71–91.

Dirmeyer, P. A., 2004: Soil moisture—Muddy prospects for a clear definition. *GEWEX News,* No. 14, International GEWEX Project Office, Silver Spring, MD, 11–12.

——, Z. Guo, and X. Gao, 2004: Comparison, validation, and transferability of eight multi-year global soil wetness products. *J. Hydrometeor.,* **5,** 1011–1033.

——, X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bull. Amer. Meteor. Soc.,* **87,** 1381–1397.

Douville, H., and F. Chauvin, 2000: Relevance of soil moisture for seasonal climate predictions: A preliminary study. *Climate Dyn.,* **16,** 719–736.

Eagleson, P. S., 1970: *Dynamic Hydrology*. McGraw-Hill, 462 pp.

Entin, J. K., A. Robock, K. Y. Vinnikov, V. Zabelin, S. Liu, A. Namkhai, and T. Adyasuren, 1999: Evaluation of Global Soil Wetness Project Soil moisture simulations. *J. Meteor. Soc. Japan,* **77,** 183–198.

Famiglietti, J. S., and E. F. Wood, 1991: Evapotranspiration and runoff from large land areas: Land surface hydrology for atmospheric general circulation models. *Surv. Geophys.,* **12,** 179–204.

Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948–present. *J. Geophys. Res.,* **113,** D01103, doi:10.1029/2007JD008470.

Fecan, F., B. Marticorena, and G. Bergametti, 1999: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas. *Ann. Geophys.,* **17,** 149–157.

Fennessey, M. J., and J. Shukla, 1999: Impact of initial soil wetness on seasonal atmospheric prediction. *J. Climate,* **12,** 3167–3180.

Friend, A. D., and N. Y. Kiang, 2005: Land surface model development for the GISS GCM: Effects of improved canopy physiology on simulated climate. *J. Climate,* **18,** 2883–2902.

Guo, Z., and P. A. Dirmeyer, 2006: Evaluation of the Second Global Soil Wetness Project soil moisture simulations. 1. Intermodel comparison. *J. Geophys. Res.,* **111,** D22S02, doi:10.1029/2006JD007233.

Hall, F. G., and Coauthors, 2006: ISLSCP Initiative II global data sets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies. *J. Geophys. Res.,* **111,** D22S01, doi:10.1029/2006JD007366.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. NCEP/Climate Prediction Center Atlas 7, Climate Prediction Center, Camp Springs, MD, 40 pp.

Holtgrieve, G. W., P. K. Jewett, and P. A. Matson, 2006: Variations in soil N cycling and trace gas emissions in wet tropical forests. *Oecologia,* **146,** 584–594.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1631–1643.

Koster, R. D., and P. C. D. Milly, 1997: The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *J. Climate,* **10,** 1578–1591.

——, and Coauthors, 2004: Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *J. Hydrometeor.,* **5,** 1049–1063.

Moorcroft, P. R., 2003: Recent advances in ecosystem-atmosphere interactions: An ecological perspective. *Proc. Roy. Soc. London,* **270,** 1215–1227.

Narasimhan, B., and R. Srinivasan, 2005: Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agric. For. Meteor.,* **133,** 69–88.

Randerson, J. T., M. V. Thompson, T. J. Conway, I. Y. Fung, and C. B. Field, 1997: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. *Global Biogeochem. Cycles,* **11,** 535–560.

Reichle, R. H., and R. D. Koster, 2004: Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.,* **31,** L19501, doi:10.1029/2004GL020938.

——, ——, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *J. Geophys. Res.,* **112,** D09108, doi:10.1029/2006JD008033.

Rodell, M., and Coauthors, 2004: The global land data assimilation system. *Bull. Amer. Meteor. Soc.,* **85,** 381–394.

——, J. Chen, H. Kato, J. S. Famiglietti, J. Nigro, and C. R. Wilson, 2007: Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeol. J.,* **15,** 159–166.

Rosenzweig, C., F. N. Tubiello, R. Goldberg, E. Mills, and J. Bloomfield, 2002: Increased crop damage in the US from excess precipitation under climate change. *Global Environ. Change,* **12,** 197–202.

Saleem, J. A., and G. D. Salvucci, 2002: Comparison of soil wetness indices for inducing functional similarity of hydrologic response across sites in Illinois. *J. Hydrometeor.,* **3,** 80–91.

Schaake, J. C., and Coauthors, 2004: An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS). *J. Geophys. Res.,* **109,** D01S90, doi:10.1029/2002JD003309.

Simmons, A. J., and J. K. Gibson, 2000: The ERA-40 Project Plan. ERA-40 Project Report Series No. 1, 63 pp. [Available online at http://www.ecmwf.int/publications/library/ecpublications/_pdf/ERA40_PRS_1.pdf.]

Thornton, P. E., and N. E. Zimmermann, 2007: An improved canopy integration scheme for a land surface model with prognostic canopy structure. *J. Climate,* **20,** 3902–3923.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.,* **107,** 4429, doi:10.1029/2001JD000659.

Yuste, J., D. D. Baldocchi, A. Gershenson, A. Goldstein, L. Misson, and S. Wong, 2007: Microbial soil respiration and its dependency on carbon inputs, soil temperature, and moisture. *Global Change Biol.,* **13,** 2018–2035.

Zhao, M., and P. Dirmeyer, 2003: Production and analysis of GSWP-2 near-surface meteorology data sets. COLA Tech. Rep. 159, 22 pp. [Available from the Center for Ocean–Land–Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705.]